

#2

520.39527X00

JC685 U.S. PRO
09/161743
01/18/01



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): SAEKI, ET AL.

Serial No.:

Filed: January 18, 2001

Title: MULTIDIMENSIONAL CROSSBAR NETWORK AND
PARALLEL COMPUTER SYSTEM

Group:

LETTER CLAIMING RIGHT OF PRIORITY

Honorable Commissioner of
Patents and Trademarks
Washington, D.C. 20231

January 18, 2001

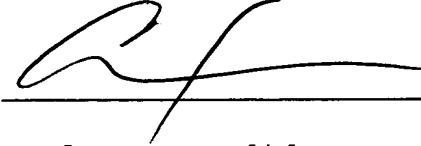
Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the
applicant(s) hereby claim(s) the right of priority based on
Japanese Patent Application No. (s) 2000-072458 filed March 10,
2000.

A certified copy of said Japanese Application is
attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Carl I. Brundidge
Registration No. 29,621

CIB/mdt
Attachment
(703) 312-6600

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

JC685 U.S. PRO
09/161743
01/18/01



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日

Date of Application:

2000年 3月10日

出願番号

Application Number:

特願2000-072458

出願人

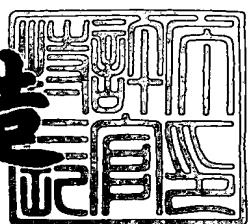
Applicant(s):

株式会社日立製作所

2000年10月20日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2000-3086898

【書類名】 特許願
【整理番号】 PNT990954
【提出日】 平成12年 3月10日
【あて先】 特許庁長官 殿
【国際特許分類】 G06F 15/173
【発明者】
【住所又は居所】 神奈川県秦野市堀山下1番地 株式会社日立製作所 工
ンタープライズサーバ事業部内
【氏名】 佐伯 裕治
【発明者】
【住所又は居所】 神奈川県秦野市堀山下1番地 株式会社日立製作所 工
ンタープライズサーバ事業部内
【氏名】 中畠 昌也
【特許出願人】
【識別番号】 000005108
【氏名又は名称】 株式会社日立製作所
【代理人】
【識別番号】 100068504
【弁理士】
【氏名又は名称】 小川 勝男
【電話番号】 03-3661-0071
【選任した代理人】
【識別番号】 100086656
【弁理士】
【氏名又は名称】 田中 恭助
【電話番号】 03-3661-0071
【手数料の表示】
【予納台帳番号】 081423
【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1
【物件名】 図面 1
【物件名】 要約書 1
【ブルーフの要否】 要

【書類名】明細書

【発明の名称】多次元クロスバーネットワークおよび並列計算機システム

【特許請求の範囲】

【請求項1】

論理的に多次元配列された複数の演算ノードを複数のクロスバースイッチによって相互接続した多次元クロスバーネットワークにおいて、

第1のクロスバースイッチと第2のクロスバースイッチとを接続する各パケット伝送路上に、上記第1、第2のクロスバースイッチと第3のクロスバースイッチに接続されるスイッチング装置を備え、

上記スイッチング装置によって、上記第1、第2、第3のクロスバースイッチ間のパケット交換を行うと共に、上記何れかのクロスバースイッチと光信号でパケット通信するためのインターフェース変換を行うことを特徴とする多次元クロスバーネットワーク。

【請求項2】

論理的に多次元配列された複数の演算ノードを相互接続するための複数のX、Y、Z軸クロスバースイッチからなる多次元クロスバーネットワークにおいて、

X軸クロスバースイッチとY軸クロスバースイッチとの間の各パケット伝送路上に、上記X軸およびY軸クロスバースイッチからの受信パケットをZ軸クロスバースイッチに選択的に転送するためのスイッチング装置を有することを特徴とする多次元クロスバーネットワーク。

【請求項3】

前記スイッチング装置が、前記X軸クロスバースイッチの入出力ポートのうちの1つに接続するための第1入出力ポートと、前記Y軸クロスバースイッチの入出力ポートのうちの1つに接続するための第2入出力ポートと、発光素子と受光素子からなる光モジュールに接続するための第3入出力ポートと、上記第1、第2入出力ポートと上記第3入出力ポートの間で送受信パケットのインターフェース変換を行うための手段とを備え、

前記Y軸クロスバースイッチとZ軸クロスバースイッチとの間のパケット伝送路の一部が上記光モジュールに結合される光ファイバによって構成されることを

特徴とする請求項2に記載の多次元クロスバーネットワーク。

【請求項4】

前記インターフェース変換手段が、前記第1、第2入出力ポートで送受信されるパケットと前記第3入出力ポートで送受信されるパケットとを互いに独立した異なる同期クロックによって処理することを特徴とする請求項2または請求項3に記載の多次元クロスバーネットワーク。

【請求項5】

それぞれ多次元配列された複数の演算ノードを複数のX、Y、Z軸クロスバースイッチによって相互接続した第1、第2のクロスバーネットワークからなる多次元クロスバーネットワークにおいて、

上記各クロスバーネットワークが、それぞれ3次元座標系で同一のY、Z座標値をもつ複数の演算ノード間でX軸方向のパケット交換を行う複数のX軸クロスバースイッチと、3次元座標系で同一Z軸座標値をもつ演算ノードを収容している複数のX軸クロスバースイッチ間でY軸方向のパケット交換を行う複数のY軸クロスバースイッチ群と、上記複数のY軸クロスバースイッチ群の間でZ軸方向のパケット交換を行う複数のZ軸クロスバースイッチとからなり、

上記第1、第2のクロスバーネットワークにおける互いに対応した位置関係にある2つのY軸クロスバースイッチを該Y軸クロスバースイッチとZ軸クロスバースイッチとの間の各パケット経路上に配置された複数のスイッチングLSIによって結合し、上記各スイッチングLSIによって上記第1、第2のクロスバーネットワーク間のパケット交換を行うようにしたことを特徴とする多次元クロスバーネットワーク。

【請求項6】

前記各スイッチングLSIが、前記2つのY軸クロスバースイッチに接続するための第1、第2入出力ポートと、発光素子と受光素子からなる第1、第2の光モジュールに接続するための第3、第4の入出力ポートと、上記第1、第2の入出力ポートからの受信パケットをヘッダ情報に従って上記第1、第2の入出力ポートのうちの他方、上記第3または第4の入出力ポートに選択的に出力すると共に、上記第3、第4の入出力ポートからの受信パケットをそれぞれ上記第1、第

2の入出力ポートに転送するための手段とからなり、

前記各Y軸クロスバースイッチとZ軸クロスバースイッチとの間のパケット伝送路の一部を上記各光モジュールに結合される光ファイバによって構成したことを特徴とする請求項5に記載の多次元クロスバーネットワーク。

【請求項7】

それぞれ3次元座標系で同一のY、Z座標値をもつ複数の演算ノード間でX軸方向のパケット交換を行うための複数のX軸クロスバースイッチと、それぞれ3次元座標系で同一Z軸座標値をもつ演算ノードを収容している複数のX軸クロスバースイッチ間でY軸方向のパケット交換を行うための複数のY軸クロスバースイッチ群と、上記複数のY軸クロスバースイッチ群の間でZ軸方向のパケット交換を行う複数のZ軸クロスバースイッチング手段とからなる多次元クロスバーネットワークにおいて、上記各Z軸クロスバースイッチング手段が、

3次元座標系で同一のX軸座標値をもつ複数のY軸クロスバースイッチの互いに対応したX軸座標位置にある入出力ポートに接続される第1入出力ポート群と、それぞれ発光素子と受光素子とを含む複数の光モジュールに接続される第2入出力ポート群とを有する第1のスイッチングLSIと、

3次元座標系で同一のX軸座標値をもつ他の複数のY軸クロスバースイッチの互いに対応したX軸座標位置にある入出力ポートに接続される第1入出力ポート群と、それぞれ発光素子と受光素子とを含む複数の光モジュールに接続される第2入出力ポート群とを有する第2のスイッチングLSIと、

上記第1のスイッチングLSIの第2入出力ポート群と上記第2のスイッチングLSIの第2入出力ポート群との間にそれぞれ光モジュールを介して結合される複数対の光ファイバとからなることを特徴とする多次元クロスバーネットワーク。

【請求項8】

それぞれL個の演算ノード収容し、L個の外部入出力ポートをもつX軸クロスバースイッチが搭載された複数のX軸ボードと、

M個の入出力ポートをもつY軸クロスバースイッチと、上記X軸クロスバースイッチに接続するための外部入出力ポートと光信号送受信用光モジュールに接続

するための入出力ポートを有し、上記Y軸クロスバースイッチの各入出力ポートに接続される複数のスイッチングLSIとが搭載された複数のY軸ボードと、

N個の入出力ポートをもつZ軸クロスバースイッチと、光信号送受信用光モジュールに接続するための入出力ポートを有し、上記Z軸クロスバースイッチの各入出力ポートに接続される複数のインターフェース変換LSIとが搭載された複数のZ軸ボードと、

上記X軸クロスバースイッチの外部入出力ポートと上記Y軸ボード上の外部入出力ポートとの間を接続する複数の電気信号線と、

上記Y軸ボード上の光信号送受信用光モジュールと上記Z軸ボード上の光信号送受信用光モジュールとの間を接続する光信号線と

からなることを特徴としたL×M×Nの多次元配列並列計算機システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、多次元クロスバーネットワークおよび多次元配列並列計算機システムに関し、更に詳しくは、多数の演算装置間を接続するためのネットワーク構成に関する。

【0002】

【従来の技術】

膨大な計算処理を高速に実行するために、例えば、複数の演算装置を密結合したマルチプロセッサ型計算機を1つの演算ノードとして、多数のノード間をネットワークで相互接続した構成の並列計算機システムが知られている。

並列計算機システムでは、多次元配列された各演算ノード毎にメモリを配置しておき、各演算ノードが、計算に必要なデータの一部を他ノードのメモリから自ノード内のメモリにコピーする分散メモリ型のものが多い。

上記分散メモリ型の並列計算機上で動作させるためにメモリ共有モデルから書き換えられたプログラムは、全ノードが同一プログラム／異データで動作するモデルに該当し、演算動作とノード間通信の現れ方が、例えば、通信中の各ノード対の一方のノードからみて、相手ノードの相対位置と通信データ量とが各ノード

対で一定となる等方転送になることが多い。メモリ共有モデルにおける最も基本的なデータアクセスは、メモリ空間の連続したアドレスでデータをアクセスするパターンを有し、上記基本パターンを分散メモリに素直に割り当てるとき、隣接ノード間でのデータ転送（隣接転送）が最も頻繁に現れる。

【0003】

この種の並列計算機システムでは、任意の2つのノード間で通信できるように演算ノードネットワークを構成する必要がある。現在では、例えば、メッシュ型、トーラス型、多次元クロスバー型、完全クロスバー型などと呼ばれるネットワーク構成によって、数百台～数千台の演算ノードからなる並列計算機システムが実現可能となっている。

「メッシュ型」および「トーラス型」のネットワーク構成では、各演算ノードに多次元論理座標を割り当てたとき、座標空間で互いに隣接した位置関係にある演算ノード間にのみネットワーク配線が存在する。従って、多数の演算ノードを論理座標と対応がとれた形で物理的に配置できれば、短い配線長で非常に多数の演算ノードを接続できる利点がある。しかしながら、このタイプのネットワーク構成では、上述した隣接転送は経路競合なしに行えるが、非隣接ノードのデータを頻繁にアクセスする等方転送の計算処理では、ノード間通信中に経路競合が発生し、システムの実効性能が低下する。従って、メッシュ型、トーラス型構成の並列計算機は、特定用途の大規模計算に適しており、汎用性には欠けるという性質がある。

【0004】

「完全クロスバー型」のネットワーク構成は、任意に組み合された複数対のノード間で同時にパケット通信できるため、各演算ノードに分散して存在するデータをグローバルにアクセスする等方転送の計算処理でも、ノード間通信ロスが少なく、効率的に計算処理を実行できる。完全クロスバー型のネットワークでは、クロスバースイッチを構成するLSI群を一箇所に集中して配置した場合、各演算ノードとクロスバースイッチとの接続配線長が長距離化する。

【0005】

「多次元クロスバー型」のネットワークは、上記2つの形式の中間的な論理構

造をもつもので、ノード間のグローバルな通信を比較的効率的に実行でき、接続ノード数とシステム性能のスケーラブルな拡張を可能とする。

多次元クロスバーネットワークは、例えば、次元数が3の場合、 $L \times M \times N$ の直方体の各格子点に演算ノードをマップしておき、X軸方向に並ぶL個のノード間を接続するクロスバースイッチと、Y軸方向に並ぶM個のノード間を接続するクロスバースイッチと、Z軸方向に並ぶN個のノード間を接続するクロスバースイッチとによって多段クロスバースイッチを構成し、演算ノード毎に、ノード内メモリと、各ノードを上記X、Y、Z軸方向の各クロスバーに選択的に接続するためのExchangerと呼ばれるスイッチとを設けた論理構成となっている。

しかしながら、多次元クロスバーネットワークでは、各ノード毎に設けられるExchangerに、X、Y、Z軸方向の各クロスバースイッチと選択的に接続するための3系統の配線が必要となるため、何れかの系統が長距離集中配線となってしまう。

【0006】

例えば、X軸クロスバースイッチと該X軸クロスバースイッチに接続されるL個の演算ノードを同一の基板上に搭載し、それぞれ同一Z座標のノード群を含むM枚の基板を同一のバックボードに収容し、該バックボード上に搭載されたL個のY軸クロスバーによって、同一X座標をもつM個のノード間を接続することによって、2次元配列のノード群までは比較的近接した配線によってネットワーク化できる。N個の2次元ノード配列からなる $L \times M \times N$ の3次元ノード配列は、N枚のバックボードに分散配置された合計 $L \times M \times N$ 個のノードを $L \times M$ 系統のZ軸クロスバースイッチの何れかのポートに接続することによって得られる。

【0007】

多次元クロスバー型は、完全クロスバー型に比較して、各軸方向のクロスバースイッチを構成するLSIの接続ポート数が少なくて済むため、設計が容易になる。また、クロスバースイッチ用のLSI群と演算ノード群を幾つかのボードに分散配置し、各ボード上で部分的に配線を済ませることによって、完全クロスバー型のような演算ノードからクロスバースイッチ部への配線の集中を回避できる。但し、上述した $L \times M$ 系統のZ軸クロスバースイッチ群を一箇所に集めて配置

した場合、バックボード上の各演算ノードとZ軸クロスバースイッチとの接続に長距離配線が必要となる。

【0008】

上記多次元クロスバー型を改良したネットワーク構成として、X、Y、Zの各軸方向への配線源となるExchanger機能を分割し、演算ノードとX軸クロスバースイッチとの間の乗換え機能は各演算ノードにもたせ、X軸クロスバースイッチとY軸クロスバースイッチとの間の乗換え機能はX軸クロスバースイッチに、Y軸クロスバースイッチとZ軸クロスバースイッチとの間の乗換え機能はY軸クロスバースイッチにもたせた構成が「分散Exchanger方式」の多次元クロスバーネットワークとして知られている。分散Exchanger方式によれば、演算ノードからX軸、Y軸、Z軸クロスバースイッチと順次に辿って、配線を一箇所に収束する接続形態でネットワークを構成できる。

【0009】

【発明が解決しようとする課題】

しかしながら、分散Exchanger方式の多次元クロスバーネットワーク構成では、Y軸クロスバースイッチとZ軸クロスバースイッチとの間の接続配線が、システム規模に応じて長距離化するため、接続ケーブルの電気的特性によって、接続可能な演算ノード数やシステム性能が制限される。

また、分散Exchanger方式の多次元クロスバーネットワークでは、最上位軸（3次元の場合はZ軸）以外の各軸方向のクロスバースイッチにおいて、別軸方向のクロスバースイッチへの乗換え機能をもたせるために、各演算ノードがExchanger機能をもつ通常の多次元クロスバーネットワークのクロスバースイッチに比較して、約2倍の外部接続ピン（LSIピン）が必要となる。

【0010】

例えば、通常の多次元クロスバーネットワーク構成では、 $L \times M \times N$ 個のノードを接続する場合、Y軸クロスバースイッチには、Y軸方向のノード数に等しいMセットの入出力ポートがあればよい。これに対して、分散Exchanger方式のY軸クロスバースイッチには、X軸クロスバースイッチに接続するためのMセットの入出力ポートと、Z軸クロスバースイッチに接続するためのMセットの入出力

ポートが必要となるため、LSIの外部接続ピン数を同一にした場合、Y軸方向に配列できるノード数が半分になってしまう。

また、分散Exchanger方式のクロスバーネットワークでは、Z軸方向に隣接する2つのノード間で通信する場合でも、X軸クロスバーとY軸クロスバーを経由しなければZ軸クロスバースイッチに到達できない。従って、最も頻繁に現れる隣接転送のノード間通信において、隣接ノードの位置関係によっては、送信パケットが宛先ノードに到達する迄の所要時間が長くなるという問題がある。

【0011】

本発明の目的は、ノード間接続コストの上昇を抑えてシステム規模（接続可能ノード数）を拡張できる分散Exchanger方式の多次元クロスバーネットワークおよび並列計算機システムを提供することにある。

本発明の他の目的は、長距離配線区間における特性劣化の少ない分散Exchanger方式の多次元クロスバーネットワークおよび並列計算機システムを提供することにある。

本発明の更に他の目的は、クロスバースイッチ当たりの収容ノード数を増加できる分散Exchanger方式の多次元クロスバーネットワークおよび並列計算機システムを提供することにある。

【0012】

【課題を解決するための手段】

上記目的を達成するために、本発明は、論理的に多次元配列された複数の演算ノードを複数のクロスバースイッチによって相互接続した多次元クロスバーネットワークにおいて、第1のクロスバースイッチと第2のクロスバースイッチとを接続するパケット伝送路上に、上記第1、第2のクロスバースイッチと第3のクロスバースイッチに接続されるスイッチング装置を備え、上記スイッチング装置によって、上記第1、第2、第3のクロスバースイッチ間のパケット交換を行うと共に、上記何れかのクロスバースイッチと光信号でパケット通信するためのインターフェース変換を行うことを特徴とする。

本発明によれば、クロスバースイッチ間を接続するスイッチング装置（LSI）に、光信号でパケット通信するためのインターフェース変換機能をもたせたこと

によって、距離の長い配線区間では光信号によってパケット通信できる。この場合、電気信号ケーブル接続ポートで送受信されるパケットと、光信号ケーブル（光ファイバ）接続ポートで送受信されるパケットとを互いに独立した異なる同期クロックによって処理することによって、長距離配線区間で光信号伝送の適した通信形態を採用できる。

【0013】

本発明において、上記スイッチング装置は、例えば、座標軸の異なるクロスバースイッチ間で受信パケットを他の座標軸のクロスバースイッチに転送するパケット分岐用スイッチとして適用される。本発明の1実施例では、論理的に多次元配列された演算ノード群を相互接続するための複数のX、Y、Z軸クロスバースイッチからなる多次元クロスバーネットワークおよび並列計算機システムにおいて、X軸クロスバースイッチとY軸クロスバースイッチとの間の各パケット伝送路上に、上記X軸およびY軸クロスバースイッチからの受信パケットをZ軸クロスバースイッチに選択的に転送するためのスイッチング装置を有することを特徴とする。

上記構成によれば、Y軸クロスバースイッチの外部でZ軸クロスバースイッチへの乗り換えが実行されるため、Y軸クロスバースイッチLSIには、Z軸クロスバースイッチとの接続ピンが不要となり、これらの接続ピンをX軸クロスバースイッチとの接続用として有効に利用できる。また、Z軸方向に隣接するノード間で通信を行う場合に、X軸クロスバースイッチからY軸クロスバースイッチを経由せずにZ軸クロスバーにパケットを転送できるため、通信パケットが宛先ノードに到達する迄の所要時間を短縮できる。

【0014】

本発明において、上記スイッチング装置は、例えば、異なる多次元クロスバーネットワーク間でのパケット交換用スイッチとしても適用される。本発明の第2実施例では、それぞれ多次元配列された複数の演算ノードを複数のX、Y、Z軸クロスバースイッチによって相互接続した第1、第2のクロスバーネットワークからなる多次元クロスバーネットワークにおいて、上記各クロスバーネットワークが、それぞれ3次元座標系で同一のY、Z座標値をもつ複数の演算ノード間で

X軸方向のパケット交換を行う複数のX軸クロスバースイッチと、3次元座標系で同一Z軸座標値をもつ演算ノードを収容している複数のX軸クロスバースイッチ間でY軸方向のパケット交換を行う複数のY軸クロスバースイッチ群と、上記複数のY軸クロスバースイッチ群の間でZ軸方向のパケット交換を行う複数のZ軸クロスバースイッチとからなり、上記第1、第2のクロスバーネットワークにおける互いに対応した位置関係にある2つのY軸クロスバースイッチを該Y軸クロスバースイッチとZ軸クロスバースイッチとの間の各パケット経路上に配置された複数のスイッチングLSIによって結合し、上記各スイッチングLSIによって上記第1、第2のクロスバーネットワーク間のパケット交換を行うようにしたことを特徴とする。

【0015】

この場合、スイッチングLSIは、具体的には、上記2つのY軸クロスバースイッチに接続するための第1、第2入出力ポートと、発光素子と受光素子からなる第1、第2の光モジュールに接続するための第3、第4の入出力ポートと、上記第1、第2の入出力ポートからの受信パケットをヘッダ情報に従って上記第1、第2の入出力ポートのうちの他方、上記第3または第4の入出力ポートに選択的に出力すると共に、上記第3、第4の入出力ポートからの受信パケットをそれぞれ上記第1、第2の入出力ポートに転送するための手段とからなり、上記各Y軸クロスバースイッチとZ軸クロスバースイッチとの間のパケット伝送路の一部が上記各光モジュールに結合される光ファイバによって構成される。

【0016】

本発明において、上記スイッチング装置は、例えば、多次元クロスバーネットワークにおける特定軸のクロスバースイッチに代わるパケット交換用スイッチとしても適用される。本発明の第3実施例では、それぞれ3次元座標系で同一のY、Z座標値をもつ複数の演算ノード間でX軸方向のパケット交換を行うための複数のX軸クロスバースイッチと、それぞれ3次元座標系で同一Z軸座標値をもつ演算ノードを収容している複数のX軸クロスバースイッチ間でY軸方向のパケット交換を行うための複数のY軸クロスバースイッチ群と、上記複数のY軸クロスバースイッチ群の間でZ軸方向のパケット交換を行う複数のZ軸クロスバースイ

ッティング手段とからなる多次元クロスバーネットワークにおいて、上記各Z軸クロスバースイッチング手段が、3次元座標系で同一のX軸座標値をもつ複数のY軸クロスバースイッチの互いに対応したX軸座標位置にある入出力ポートに接続される第1入出力ポート群と、それぞれ発光素子と受光素子とを含む複数の光モジュールに接続される第2入出力ポート群とを有する第1のスイッチングLSIと、3次元座標系で同一のX軸座標値をもつ他の複数のY軸クロスバースイッチの互いに対応したX軸座標位置にある入出力ポートに接続される第1入出力ポート群と、それぞれ発光素子と受光素子とを含む複数の光モジュールに接続される第2入出力ポート群とを有する第2のスイッチングLSIと、上記第1のスイッチングLSIの第2入出力ポート群と上記第2のスイッチングLSIの第2入出力ポート群との間にそれぞれ光モジュールを介して結合される複数対の光ファイバとからなることを特徴とする。上記構成によれば、ノード間接続コストを抑えて3次元クロスバーネットワークを構成できる。

【0017】

【発明の実施の形態】

以下、本発明の実施例を図面を参照して説明する。

図1は、本発明による分散Exchanger方式の3次元クロスバーネットワークにおけるX、Y、Z軸方向のクロスバースイッチ間の接続関係を説明するための図である。

図において、20はX軸クロスバースイッチ(LSI)21を搭載したX軸ボード、30はY軸クロスバースイッチ(LSI)31を搭載したY軸ボード、40はZ軸クロスバースイッチ(LSI)41を搭載したZ軸ボードを示す。

【0018】

$L \times M \times N$ 個のノードを接続する3次元クロスバーネットワークを構成する場合、X軸ボード20には、X軸上の座標値と対応するL個のノードボード10が収容される。22-1～22-Lは、これらのノードボードを接続するためのコネクタを示す。各コネクタ22は、X軸ボードの表面あるいは内部に形成されたプリント配線によってX軸クロスバースイッチ21の入出力ポートと結合されている。

【0019】

各ノードボード10は、図2に示すように、複数のLSIからなる演算処理部11と、ネットワークインターフェースLSI:12と、これらのLSIで共有して使用されるメモリLSI:13と、メモリスイッチ14とを有し、これらの回路要素によって演算ノードが形成されている。15は、ボード端部に設けられたリード端子群であり、これらのリード端子をX軸ボード20上のコネクタ22-iに挿し込むことによって、演算ノードがX軸クロスバースイッチ21の所定の入出力ポートに接続される。

演算処理部11は、他の演算ノードにデータを送信するとき、メモリ13の所定の領域に送信データと制御情報を設定した後、ネットワークインターフェースLSI:12を起動する。起動されたネットワークインターフェースLSI:12は、メモリ13に用意された制御情報に従って送信データを読み出し、例えば、8バイト幅のパケットを生成して、X軸クロスバースイッチ21に送出する。上記パケットのヘッダ部には、宛先アドレスとして、3次元クロスバーネットワークにおける宛先ノード位置を示すX、Y、Z座標値が設定される。

【0020】

以下、宛先アドレス[i, j, k]をもつパケットに着目して、図1の3次元クロスバーネットワーク上での転送動作を説明する。

X軸クロスバースイッチ21は、上述したノードボード（演算ノード）10を収容するためのL個の入出力ポートの他に、Y軸クロスバースイッチ31との接続線を収容するためのL個の外部入出力ポートを備える。X軸クロスバースイッチ21とY軸クロスバースイッチ31との接続には、例えば、同軸ケーブルからなる入出力線200[1, y, z]～200[L, y, z]が利用される。

【0021】

ここで、鍵括弧内に文字は、3次元クロスバーネットワークにおける演算ノードのアドレス（X、Y、Z座標値）を示しており、例えば、入出力線200[1, y, z]には、コネクタ22-1に接続されたアドレス[1, y, z]をもつ演算ノードから出力された他の演算ノード宛の送信パケット、あるいは他の演算ノードから上記アドレス[1, y, z]をもつ演算ノード宛の受信パケットが流

れることを意味している。また、座標値 [y, z] は、X軸クロスバースイッチ21に収容された演算ノード群に共通するY、Z座標値であり、X軸クロスバースイッチ21には、上記座標値 [y, z] がボードアドレスとして予め設定されている。

【0022】

X軸クロスバースイッチ21は、収容している何れかの演算ノードから宛先アドレス [i, j, k] のパケットを受信すると、パケットヘッダの宛先X座標に従って、第iポートにスイッチングする。また、受信パケットの宛先Y、Z座標j、kを上記X軸ボードアドレスと比較し、Y軸上またはZ軸上でのスイッチングを要する受信パケットについては、上記第iポートの外部入出力線、例えば、200 [i, y, z] に送出する (i = 宛先X座標値)。宛先Y、Z座標値j、kがボードアドレス [y, z] と一致し、Y軸、Z軸上でのスイッチングを必要としないパケットについては、第iポートのノードボード側 (コネクタ22-i) に送出する。

【0023】

Y軸ボード30は、図3、図4で後述するように、3次元クロスバーネットワークのZ座標値毎に、X座標と対応してL個ずつ用意され、それぞれボードアドレスとして特定のX、Z座標値が割り当てられている。各Y軸ボードには、Y軸クロスバースイッチ31と、3次元クロスバーネットワークのY座標と対応したM個のスイッチングLSI : 31-1 ~ 31-Mとが搭載されている。

本実施例において、各スイッチングLSI : 31は、X軸、Y軸のクロスバースイッチ間でZ軸クロスバースイッチ方向にパケットを分岐し、分岐パケットについて光伝送のためのインターフェース変換を行うためのものであり、以下の説明では、インターフェース変換LSIと呼ぶことにする。

【0024】

X軸クロスバースイッチ21から入出力線200 [i, y, z] に送出されたパケットは、X、Z座標値 [i, z] をボードアドレスとするY軸ボード30の第y番目のインターフェース変換LSI : 31-yに入力される。上記インターフェース変換LSI : 31-yは、入出力配線200 [i, y, z] からの受信パケ

ットの宛先アドレスをチェックし、Y軸上でのスイッチングを必要としないパケット、即ち、宛先Y座標値jが上記インターフェース変換LSI:31-yのY座標yに一致したパケットは、レーザー発光素子33T-yと受光素子33R-yとからなる光モジュール33-yに接続された入出力線320-y(=320-j)に送出し、Y軸上でのスイッチングを要するパケット、すなわち、宛先Y座標値jがインターフェース変換LSIのY座標値yに一致しないパケットは、Y軸クロスバースイッチ31に接続された入出力線310-yに出力する。

【0025】

入出力線310-yに出力されたパケットは、Y軸クロスバースイッチ31によって、宛先Y座標値jと対応した第j番目のインターフェース変換LSI:31-jにスイッチングされる。上記インターフェース変換LSI:31-jは、Y軸クロスバースイッチ31からの受信パケットについて宛先アドレスをチェックし、Z軸上でのスイッチングを必要としないパケット、すなわち、宛先Z座標値kがY軸ボードアドレスのZ座標値zと一致したパケットについては、入出力線200[i, j, z]に出力し、Z軸上でのスイッチングを要するパケットは、入出力回線320-jに出力する。

入出力線320-jに出力されたパケットは、レーザー発光素子33T-jで光信号に変換され後、入出力線(光ファイバ)300[i, j, z]に送出される。上記入出力線(光ファイバ)300[i, j, z]は、ボードアドレスとしてX、Y座標値[i, j]をもつZ軸ボード40に接続されている。

【0026】

各Z軸ボード40は、Z軸クロスバースイッチ41と、3次元クロスバーネットワークのZ座標値と対応したN個のインターフェース変換LSI:42-1~42-Nとを有し、各インターフェース変換LSI:42-q(q=1~N)は、Z軸クロスバースイッチ41に接続された入出力線410-qと、レーザー発光素子43T-qと受光素子43R-qとを含む光モジュール43-qに接続された入出力線420-qとを備えている。

【0027】

Y軸ボード30上のインターフェース変換LSI:32とは異なり、Z軸ボード

40上の各インターフェース変換LSI:42の入出力ポートは、Z軸クロスバースイッチ側と光モジュール側の2つとなっている。このため、Y軸ボード30から入出力線（光ファイバ）300[i, j, z]に光信号として送出されたパケットは、Z軸ボード40側の受光素子44R-zで電気信号に変換された後、無条件で入出力線410-zに転送され、Z軸クロスバースイッチ41に入力される。Z軸クロスバースイッチ41は、上記パケットを宛先Z座標値kに対応した第kインターフェース変換LSI:42-kにスイッチングする。

上記第kインターフェース変換LSI:42-kは、Z軸クロスバースイッチ41からの入力パケットを入出力線420-kに転送する。従って、上記パケットは、レーザー発光素子43T-kで光信号に変換された後、光ファイバ300[i, j, k]に送出される。

【0028】

上記光ファイバ300[i, j, k]は、ボードアドレスとしてX、Z座標値[i, k]をもつY軸ボード30の第jインターフェース変換LSI:32-jの光モジュールに接続されている。この時点では、受信パケットの宛先X、Z座標は、上記Y軸ボードのボードアドレス[i, k]と既に一致しており、且つ、宛先Y座標jが上記第jインターフェース変換LSI:32-jのアドレスjと一致しているため、Y軸クロスバースイッチ31でスイッチングする必要はない。このため、上記受信パケットは、上記Y軸ボード上の第jインターフェース変換LSI:32-jによって入出力線200[i, j, k]に出力される。

上記入出力線200[i, j, k]は、ボードアドレスとしてY、Z座標値[i, k]をもつX軸ボード20のX軸クロスバースイッチ21に接続されている。上記X軸クロスバースイッチ21は、入出力配線200[i, j, k]からの上記パケットを受信すると、これを第iコネクタ22-iが接続された第iポートに送出する。これによって、上記パケットは、パケットヘッダの宛先アドレス[i, j, k]に対応した演算ノードで受信処理されることになる。

【0029】

図3は、同一のZ座標値zをもつX軸クロスバースイッチ群とY軸クロスバースイッチ群との接続関係を示す。ここでは、図面の簡略化のために、各インタフ

エース変換LSIに接続される光モジュールは省略してある。

$L \times M \times N$ 個の演算ノードからなる3次元クロスバーネットワークの各XY平面を構成する2次元クロスバーネットワークは、Y座標値と対応して用意されたM個のX軸ボード $20-1 \sim 20-M$ と、X座標値と対応して用意されたL個のY軸ボード $30-1 \sim 30-L$ とからなる。X軸ボード $20-1 \sim 20-M$ 上のM個のX軸クロスバースイッチ $21-1 \sim 21-M$ で形成される $L \times M$ 個の外部入出力ポートのうち、同一のX座標値 i をもつM個の入出力ポートが、入出力線 $200[i, 1, z] \sim 200[i, M, z]$ を介して、ボードアドレス $[i, z]$ をもつY軸ボード $30-i$ 上のクロスバースイッチ $31-i$ に接続される。

【0030】

Y軸ボード $30-1 \sim 30-L$ 上の各インターフェース変換LSIには、夫々に接続された入出力線 $200[i, 1, z] \sim 200[i, M, z]$ と対応したアドレスを割り当てることができる。以下の説明では、Y軸ボード $30-i$ 上の配線 $200[i, j, z]$ と接続されたインターフェース変換LSIは、符号 $32[i, j, z]$ 、上記インターフェース変換LSIの光モジュール $33, 34$ に接続された光信号入出力線（光ファイバ）は、符号 $300[i, j, z]$ で示すこととする。

【0031】

図4は、複数のXY平面を統合するZ軸ボード40とY軸ボード30との間の接続関係を示す。

図において、 $30-1-1 \sim 30-L-1$ は、Z座標値が「1」のY軸ボード群、 $30-1-2 \sim 30-L-2$ は、Z座標値が「2」のY軸ボード群を示し、 $30-1-N \sim 30-L-N$ は、Z座標値が「N」のY軸ボード群を示す。また、各Y軸ボードとZ軸ボードとの接続関係を光信号入出力線 $300[1, 1, 1] \sim 300[L, M, N]$ によって示す。

光信号入出力線 300 に付されたアドレス値から明らかなように、3次元クロスバーネットワークの構成に必要なZ軸ボード40の枚数は、X、Y座標値の組み合わせ数に対応しており、各Z軸ボードは、N層のXY平面から引き出された同一X、Y座標値をもつ入出力線 300 間でパケット交換できるように、それぞ

れ特定のY軸ボード群と結合されている。

【0032】

図5は、Y軸ボード30上に搭載され、Y軸クロスバースイッチ31とZ軸クロスバースイッチ41との間でのパケット交換機能を備えるインターフェース変換LSI:32の1実施例を示す。

インターフェース変換LSI:32は、X軸クロスバースイッチ21と接続する入出力線200を収容するための第1入出力ポート51Aと、Y軸クロスバースイッチ31と接続する入出力線310を収容するための第2入出力ポート51Bと、光モジュール33と接続する入出力線320を収容するための第3入出力ポート51Cとを有し、上記各入出力ポートは入力ポートINと出力ポートOUTを備える。

CLK1は、X軸ボード20とY軸ボード30間の電気信号によるデータ転送のための基準クロックを示し、CLK2は、Y軸ボード30とZ軸ボード40間の光信号によるデータ転送のための基準クロックを示す。PLL回路60と61は、それぞれ上記クロックCLK1、CLK2に同期した内部クロックを発生する。

【0033】

第1入出力ポート51Aの入力ポートINで受信されたパケット信号は、位相調整回路52Aを経てデータバッファ53Aに入力される。上記位相調整回路52Aは、PLL回路60で生成される内部クロックに基づいて入力信号の位相を調整し、入力ポートINから入力されたパケット信号を、例えば、8バイト幅のデータとして上記データバッファ53Aに蓄える。

上記第1入出力ポート51Aの出力ポートOUTには、後述するように、セレクタ55Aで選択された第2、第3入出力ポートからの入力パケットが出力される。尚、上記セレクタ55Aの選択動作は、出力制御回路56Aによって制御されている。以下の説明では、入出力線200と対応する回路要素51A～56Aからなる回路部分を第1インターフェースユニットと呼び、入出力線310、320と対応する同様の回路部分を第2、第3インターフェースユニットと呼ぶことにする。

【0034】

入力パケット制御回路54Aは、位相調整回路52Aの出力を監視し、パケットの先頭を示すスタートフラグを識別して、パケットヘッダの所定位置に設定された宛先Y座標を抽出する。上記入力パケット制御回路54Aは、各パケットから抽出した宛先Y座標値と、予め記憶されている各インターフェース変換LSI:30に固有のY座標値とを比較し、Y座標値が一致した場合は、第3インターフェースユニットの出力制御回路56Cに、また、Y座標値が一致しない場合は、第2インターフェースユニットの出力制御回路56Bに対して、それぞれパケットの出力許可要求信号REQ1を送り、出力制御回路56Bまたは56Cが返送するパケット出力許可信号C20またはC30に応答して、データバッファ53A内のパケットを内部バス62Aに出力する。上記内部バス62Aは、第2インターフェースユニットの出力セレクタ55Bと、第3インターフェースユニットの出力セレクタ55Cに接続されている。

【0035】

第2インターフェースユニットも、上記第1インターフェースユニットと同様に、位相調整回路52B、データバッファ53B、入力パケット制御回路54B、セレクタ55B、出力制御回路56Bを備えている。

第2インターフェースユニットの入力ポートINには、Y軸クロスバースイッチ31によってY軸方向のスイッチングを終えたパケットが入力されるため、入力パケット制御回路54Bでは、入力パケットのヘッダから宛先Z座標値を抽出し、これを予め記憶されているボードアドレスのZ座標値とを比較する。

Z座標値が一致した場合は、第1インターフェースユニットの出力制御回路56Aに、また、Z座標値が一致しない場合は、第2インターフェースユニットの出力制御回路56Bにパケットの出力許可要求信号REQ2を送り、出力制御回路56Aまたは56Cが返送するパケット出力許可信号C10またはC30に応答して、データバッファ53B内のパケットを内部バス62Bに出力する。上記内部バス62Bは、第1インターフェースユニットの出力セレクタ55Aと、第3インターフェースユニットの出力セレクタ55Cに接続されている。

【0036】

第3インターフェースユニットは、データバッファ53C、入力パケット制御回路54C、セレクタ55C、出力制御回路56Cの他に、同期化回路57、58と、データ幅変換回路59を備えている。出力制御回路56Cは、要求信号REQ1またはREQ2を受信すると、要求発生順にパケット出力許可信号C30を返送し、セレクタ55Cを切替えることによって、要求元と対応する内部バス62Aまたは62Bのパケットデータを同期化回路57に取り込む。

同期化回路57は、PLL回路61が発生する内部クロックに同期して動作し、第3出入力ポート51Cにおけるデータ送信を外部クロックCLK2に同期させる。上記同期化回路57から出力された8バイト幅のパケットは、データ幅変換回路59によって、例えば、22ビット幅のパケットデータに変換された後、第1、第2出入力ポートにおけるデータ転送の4倍の速度（周波数）で出力ポートOUTから送出される。出力ポートOUTから送出されたパケットデータは、レーザ発光素子33Tで光信号に変換され、光ファイバ300を介してZ軸ポート40に送られる。尚、上記出入力ポート51Cと、レーザ発光素子33Tおよび受光素子33Rを含む光モジュール33との間では、上記パケットデータの他に同期クロックも送受信される。

【0037】

Z軸クロスバースイッチ41でZ軸方向のスイッチングを終えたパケットは、光ファイバ300を介して受光素子33Rに入力され、電気信号に変換された後、上記第3出入力ポート51Cの入力ポートINに入力される。入力ポートINで受信されたパケットデータは、データ幅変換回路51Cで8バイト幅に変換された後、同期化回路58を経て、データバッファ53Cに蓄積される。尚、同期化回路58は、PLL回路60が発生する内部クロックに同期して動作している。

【0038】

入力パケット制御回路54Cは、同期化回路58からデータバッファ53Cに供給されるパケットデータを監視し、パケットヘッダの宛先Y座標を抽出して、予め記憶されているインターフェース変換LSI:30に固有のY座標値とを比較

する。Y座標値が一致した場合は、第1インターフェースユニットの出力制御回路56Aに、また、Y座標値が一致しない場合は、第2インターフェースユニットの出力制御回路56Bに対して、それぞれパケットの出力許可要求信号REQ3を送り、出力制御回路56Aまたは56Bが返送するパケット出力許可信号C10またはC20に応答して、データバッファ53Cから内部バス62Cにパケットデータを出力する。

内部バス62Cは、第1インターフェースユニットの出力セレクタ55Aと、第2インターフェースユニットの出力セレクタ55Bに接続されている。尚、第3入出力ポート51Cで入出力されるパケットは、既にY軸方向のスイッチングを終えた状態となっているため、実際の動作において、入力パケットが第2インターフェースユニットに転送されることはない。

【0039】

第1、第2インターフェースユニットにおける出力セレクタ55A、55Bも、セレクタ55Cと同様に、それぞれ出力制御回路56A、56Bによって制御され、内部バス62Bと62C、62Aと62Cの中からパケット出力を許可された内部バス選択する。

上記構成によって、上記インターフェース変換LSI:32は、第1、第2、第3の3つのインターフェースユニット間で、受信パケットの宛先アドレスと該LSIの固有アドレスとの関係に応じたパケット転送動作を行うことができ、これをY軸ボードに適用することによって、図1で説明した3次元配列の任意の演算ノード間のパケット転送が可能となる。

【0040】

上記インターフェース変換LSIでは、電気信号インターフェース側で使用する同期クロックCLK1と、光信号インターフェース側で使用する同期クロックCLK2とを独立させ、X軸クロスバースイッチとY軸クロスバースイッチとの間ではクロックCLK1に同期した電気信号によってパケット転送し、Y軸ボードとZ軸ボードとの間ではクロックCLK2に同期した光信号でパケット転送するようになっているため、位相調整回路52に必要な同期クロックCLK1の分配範囲を、例えば、同軸ケーブルで分配容易な範囲に局所化することができる。

【0041】

従って、上記インターフェース変換LSIを使用すれば、同一クロック源による同期クロック分配範囲、例えば、1群のX、Yボードで構成されるL×M個の演算ノードを含むクロスバーネットワークを単位として、クロック源の異なる複数のクロスバーネットワーク間を容易に相互接続できる。また、小規模クロスバーネットワークの需要には上記L×M個の演算ノードからなる基本ユニットで対処し、必要に応じてZ軸ボードと光ファイバによる接続ユニット数を増加することによって、需要に応じたクロスバーネットワーク規模の拡張（ノード数の増設）が極めて容易になる。

【0042】

図6は、インターフェース変換LSI:32の実用的な形態の1例を示す。

ここに示した例では、図5で説明した回路構成を同一LSI基板上に2組搭載し、51A-1、51B-1、51C-1を入出力ポートとする第1の回路部と、51A-2、51B-2、51C-2を入出力ポートとする第2の第2の回路部で、PLL回路60と61を共用するようにしている。

このように、図5の回路構成を同一LSI基板上に複数組搭載し、Y軸ボード30上でのインターフェース変換LSI:32の占有面積を減らすことによって、3次元クロスバーネットワークの小型化と、接続ノード数の増加が可能となる。

【0043】

Z軸ボード40に搭載されるインターフェース変換LSI:41は、図5の回路構成において、入出力線200に接続される第1インターフェースユニットが不要となり、Z軸クロスバースイッチに接続される第2インターフェースユニットと、光モジュール33に接続される第3インターフェースユニットのみが必要となる。この場合、内部バス上での出力パケットの競合が発生しないため、各インターフェースユニットは、入力パケット制御回路54B、54C、セレクタ55B、55C、出力競合制御回路56B、56Cを省略した簡単な回路構成を採用できる。上記Z軸ボード用のインターフェース変換LSI:41についても、図6と同様、同一LSI基板上に複数組の回路部を搭載した装置構成を採用できる。

【0044】

図7は、Y軸クロスバースイッチ31の主要部の構成を示す。

Y軸クロスバースイッチ31は、入出力線310-j ($j = 1 \sim M$) と対応した複数のインターフェースユニット70-jをデータバスと制御線で相互接続した構成となっている。X軸クロスバースイッチ41もこれと同様の構造をもつ。

【0045】

各インターフェースユニット70-jは、入出力線310-jを接続するための入出力ポート51と、位相調整回路52と、データバッファ53と、入力パケット制御回路54と、セレクタ55と、出力制御回路56、クロックCLK1に同期した内部クロックを発生するPLL回路60とからなり、図5で説明したインターフェース変換LSI32における第1、第2のインターフェースユニットと類似した機能を備える。

即ち、入力パケット制御回路54は、位相調整回路からデータバッファ53に入力されるパケットデータを監視し、パケットヘッダから宛先Y座標を抽出すると、上記宛先Y座標値qに対応した第qインターフェースユニットの出力制御回路に接続された制御線72A-qに、パケット出力許可の要求信号REQを発行する。この状態で、上記第qインターフェースユニットから、制御線73A-qを介して出力許可信号CNTを受信すると、入力パケット制御回路54は、データバッファ53からデータバス71にパケットデータを出力する。

【0046】

上記データバス71は、各インターフェースユニットに専用のバスとなっており、セレクタ55には、インターフェースユニットの個数と対応した複数のバスが接続されている。上記セレクタ55を制御する出力制御回路56は、制御線72Bと73Bを介して、他の複数のインターフェースユニット70-j ($j = 1 \sim M-1$) の入力パケット制御回路54に接続されている。

出力制御回路56は、何れかの制御線72B-kから要求信号REQを受信すると、これに対応する制御線73B-kに出力許可信号CNTを返送し、制御線72B-kと対応するバス71-kが選択されるようにセレクタ55を制御する。制御線72Bから複数の要求信号REQを受信した場合、出力制御回路56は

、所定のアルゴリズム、例えば、要求信号発生順に選ばれた制御線 73B-k に対して出力許可信号 CNT を返送する。

【0047】

図8は、Y軸クロスバースイッチ31の実用的な1実施例を示す。

図1に示したY軸クロスバースイッチ31は、各インターフェース変換LSI32との間で、例えば、8バイト単位でパケットデータを送受信する。この場合、図7に示した各インターフェースユニット70-jの入出力ポート51に8バイト分の外部接続ピンが必要となる。また、各インターフェースユニット間の接続バス71に各々8バイト幅の信号線が必要となるため、LSIの構造が複雑化し、ポート数の増加が困難となる。

図8に示した実施例は、Y軸クロスバースイッチ31をLSI化した場合に必要となる外部接続ピン数を少なくし、構造を簡素化することによってLSI化を容易にしたものであり、各入出力ポートが2バイト幅でパケットデータを送受信できるように、Y軸クロスバースイッチ31をLSI:31A~31Dに4分割（バイトスライス構造）したことを特徴としている。

【0048】

Y軸クロスバースイッチ31に上記構造を採用した場合、図5に示した各インターフェース変換LSI:32の第2入出力ポート51Bにおいて、入力ポートINと出力ポートOUTに接続される入出力線310をそれぞれ2バイト幅をもつ第1~第4の入出力線に分割し、図8に示すように、上記各LSI:31A~31Dに2バイト幅で並列的に入出力するようにしておく。また、各クロスバースイッチ31が2バイト幅のパケットを受信して経路選択できるようにするため、各パケットヘッダには、宛先ノードアドレスやデータ長等の情報を予め4重化して与えておく。上記分割構造は、X軸クロスバースイッチ41も同様に適用できる。

【0049】

多数のLSIと入出力線（ケーブル）接続用外部コネクタの搭載を必要とするY軸ボード30およびZ軸ボード40は、例えば、多層プリント配線基板の表面側に、クロスバースイッチLSI:31（31A~31D）と、入出力線（例え

ば、同軸ケーブル) 200を接続するためのコネクタ群とを配置し、上記基板の裏面側に、インターフェース変換LSI群32と、それぞれ複数の光モジュール33を搭載した光ボード接続用のコネクタとを搭載し、各インターフェース変換LSI:32とクロスバースイッチLSI:31との接続(入出力線310)と、上記入出力線接続コネクタ群とインターフェース変換LSI群32との間の接続は、上記基板を貫通して一方の面から他方の面に延びて形成されるプリント配線によって達成する。

【0050】

Y軸ボードの構成に必要なその他の回路要素、例えば、上記クロスバースイッチLSI:31にデータ転送用クロック信号を分配するためのLSIは基板表面上に、上記光モジュール33にクロック信号を分配するためのLSIは基板裏面に、上記クロスバースイッチLSI:31およびインターフェース変換LSI:32に初期値を設定するために必要なLSIと、終端抵抗やノイズ低減用のコンデンサは基板の両面に配置し、上記何れかの面に電源ボード搭載用コネクタを搭載すればよい。

【0051】

図9は、本発明による分散Exchanger方式の3次元クロスバーネットワークの第2の実施例を示す。

本実施例は、図1～図4で説明した3次元クロスバーネットワークを2組使用し、Y軸ボード30とZ軸ボード40の間に、図10で後述する4つのインターフェースユニット間のスイッチング機能を備えるスイッチングLSI:80を介在させることによって、第1クロスバーネットワークから第2クロスバーネットワークに、あるいはその逆方向にパケット転送を可能としたものである。

【0052】

図9において、符号Aが付されたX軸クロスバースイッチ21-1A～21A-M、Y軸クロスバースイッチ31A、Z軸クロスバースイッチ41A-1～41A-Mは、第1クロスバーネットワークを構成し、符号Bが付されたX軸クロスバースイッチ21-1B～21B-M、Y軸クロスバースイッチ31B、Z軸クロスバースイッチ41B-1～41B-Mは、第2クロスバーネットワークを

構成している。上記Y軸クロスバースイッチ31A、31Bは、それぞれ図4に示したY軸ボード30-1-1～30-L-Nのうちの1つに相当する。

【0053】

図9において、インターフェース変換LSI:42-1～42-Mは、1つのLSI基板に2ポート分の回路を含んでおり、実質的には、Z軸クロスバースイッチ41A-1～41A-M用のLSI群と、Z軸クロスバースイッチ41B-1～41B-M用のLSI群とに分かれている。

本実施例では、X軸クロスバースイッチ群に接続される各演算ノードに、ノードアドレスとして、3次元クロスバーネットワークにおける座標値[x, y, z]と、各ノードが所属する3次元クロスバーネットワークの識別子（セット識別子）[t]を与えておき、演算ノードと対応したX軸クロスバースイッチの各入出力線200が[x, y, z, t]のアドレスをもつようとする。また、各演算ノードから送信されるパケットのヘッダには、宛先アドレスとして、宛先ノードの座標値とネットワーク識別子を設定する。ここでは、第1、第2のクロスバーネットワークの識別子tをそれぞれA、Bで表すこととする。

【0054】

Y軸クロスバースイッチ31A、31Bは、それぞれX軸クロスバースイッチに接続するための第1の入出力ポート群と、X軸クロスバースイッチに接続するための第2の入出力ポート群を備え、第1の入出力ポート群は、入出力線200によってX軸クロスバースイッチの入出力ポートに直結され、第2の入出力ポート群は、入出力線群340Aまたは340Bによって、スイッチングLSI:80-1～80-Mの第1または第2入出力ポートに接続される。

上記Y軸クロスバースイッチ31A、31Bでは、X軸クロスバースイッチから受信したパケットの宛先Y座標に従ってY軸方向のスイッチングを行い、Z軸方向またはネットワーク間スイッチングを要するパケットについては、スイッチングLSI側の出力ポートに送出し、これらのスイッチングを必要としないパケットについては、X軸クロスバースイッチ側の出力ポートに送出する。

【0055】

上記各スイッチングLSI:80-i (i=1～M)は、Y軸クロスバースイ

ツチ31Aから受信したパケットの宛先アドレスをチェックし、宛先ネットワーク識別子 t が予め上記受信パケットの入力ポートに割り当てられたネットワーク識別子と異なる場合には、ネットワーク間のスイッチングを行ない、Z軸方向のスイッチングを要するパケットは、光モジュール83側の出力ポートに送出し、Z軸方向のスイッチングを必要としないパケットは、他方のネットワークに属したX軸クロスバースイッチ側の出力ポートに送出する。上記光モジュール83は、光ファイバによってZ軸ボード上の光モジュール43に接続されているため、前述した第1実施例と同様に、Z軸クロスバースイッチ41Aまたは41BによるZ軸方向のスイッチングが可能となる。

【0056】

図10は、スイッチングLSI:80の構成を示す。

スイッチングLSI:80は、Y軸クロスバースイッチ31との間でパケットを送受信する2つのインターフェースユニット81A、81Bと、光モジュール83に接続される2つのインターフェースユニット81C、81Dとからなり、これらのインターフェースユニット間でパケット交換を行う。図9に示したクロスバーネットワークに適用した場合、インターフェースユニット81Aと81Cが第1クロスバーネットワーク用、インターフェースユニット81Bと81Dが第2クロスバーネットワーク用となる。

インターフェースユニット81Aと81Bは、図5で説明したインターフェース変換LSI:32の第1インターフェースユニットと同様の構成要素からなり、インターフェースユニット81Cと81Dは、上記インターフェース変換LSIの第3インターフェースユニットと同様の構成要素からなっている。

【0057】

スイッチングLSI:80を図9のクロスバーネットワークに適用する場合、インターフェースユニット81Aの入力パケット制御回路54Aで、入力ポートINから入力された受信パケットのヘッダから宛先Z座標値 z と宛先ネットワーク識別子 t を抽出し、予め記憶している固有のZ座標値およびネットワーク識別子（この例ではA）と比較する。

宛先Z座標値 z と宛先ネットワーク識別子 t の双方が固有値と不一致の場合は

、第2クロスバーネットワーク用の光モジュール側のインターフェースユニット81Dの出力制御回路56Dにパケット出力許可要求を出し、出力許可信号を待って、データバッファ53Aから内部バス63Aにパケットを出力する。

宛先Z座標値zが一致し、宛先ネットワーク識別子t不一致の場合は、第2クロスバーネットワーク用のX軸クロスバースイッチ側のインターフェースユニット81Bの出力制御回路56Bに、宛先Z座標値zが不一致で、宛先ネットワーク識別子tが一致した場合は、第1クロスバーネットワーク用の光モジュール側のインターフェースユニット81Cの出力制御回路56Cにパケット出力許可要求を出し、出力許可信号を待って、データバッファ53Aから内部バス63Aにパケットを出力する。

【0058】

インターフェースユニット81Bの入力パケット制御回路54Bは、固有値として入力パケット制御回路54Aとは異なるネットワーク識別子（この例ではB）を記憶している。入力パケット制御回路54Bは、受信パケットの宛先Z座標値zと宛先ネットワーク識別子tの双方が固有値と不一致の場合は、第1クロスバーネットワーク用の光モジュール側のインターフェースユニット81Cの出力制御回路56Cに、宛先Z座標値zが一致し、宛先ネットワーク識別子tが不一致の場合は、第1クロスバーネットワーク用のX軸クロスバースイッチ側のインターフェースユニット81Aの出力制御回路56Aに、宛先Z座標値zが不一致で、宛先ネットワーク識別子tが一致した場合は、第2クロスバーネットワーク用の光モジュール側のインターフェースユニット81Dの出力制御回路56Dに、パケット出力許可要求を出す。そして、出力制御回路56A、56Cまたは56Dからの出力許可信号を待って、データバッファ53Bから内部バス63Bにパケットを出力する。

【0059】

インターフェースユニット81Cと81Dの入力ポートINには、既にY軸、Z軸方向およびネットワーク間のスイッチングを完了したパケットが入力されるため、受信パケットをそれぞれY軸クロスバースイッチ側のインターフェースユニット81A、81Bに転送すればよい。従って、インターフェースユニット81Cと

81Dの入力パケット制御回路54C、54Dは、入力ポートINからパケットが入力された時、それぞれ転送先インターフェースユニットの出力制御回路56A、56Bにパケット出力許可要求を出し、出力許可信号を待って、データバッファ53C、53Dから内部バス63C、63Dにパケットを出力すればよい。しかしながら、図10に示した実施例では、スイッチングLSI:80の用途を汎用化するため、内部バス63Cと63Dをインターフェースユニット81A、81Bの双方の出力セレクタ55A、55Bに接続し、パケット転送先を選択できるようにしてある。

【0060】

インターフェースユニット81A、81Bでは、それぞれのセレクタ55A、55Bに他の3つのインターフェースユニットの内部バスを接続し、そのうちの1つを出力制御回路56A、56Bで選択することによって、3つの入力ポートからの受信パケットを出力ポートOUTに選択的に送出できるようになっている。また、インターフェースユニット81C、81Dでは、それぞれのセレクタ55C、55DにY軸クロスバースイッチ側の2つの内部バス63A、63Bを接続し、そのうちの1つを出力制御回路56C、56Dで選択することによって、2つの入力ポートからの受信パケットを出力ポートOUTに選択的に送出できるようになっている。

【0061】

上記第2実施例の構成によれば、スイッチングLSI:80を4次元目のスイッチとして利用し、Z軸クロスバースイッチを2組の3次元クロスバーネットワークで共用させることによって、ハードウェア規模を倍増することなく、演算ノード数の倍増することができる。また、上述した分散Exchanger方式クロスバースイッチ機能を併せ持つスイッチングLSI:80をZ軸クロスバースイッチ側にも適用すれば、安価な間接続コストで更に次元数の高いクロスバーネットワークを構成することが可能となる。

【0062】

図11は、本発明による分散Exchanger方式の3次元クロスバーネットワークの第3の実施例を示す。

本実施例は、図1～図4で説明した3次元クロスバーネットワークにおけるZ軸クロスバースイッチ31の機能をスイッチングLSI:80によって実現したことを特徴とする。

図において、Y軸クロスバースイッチ31A、31B、31C、31Dは、それぞれ異なるZ座標値z、例えば、z=1、2、3、4をもつ。

【0063】

本実施例では、第2実施例と同様に、Y軸クロスバースイッチ31Aと31Bの同一Y座標をもつ入出力ポート同士を第1のスイッチングLSI群80-1～80-Mにより接続し、これらのLSI群におけるパケットのスイッチング機能を利用して、Y軸クロスバースイッチ31Aと31Bの間でのZ軸方向のスイッチングを行う。これと同様に、Y軸クロスバースイッチ31Cと31Dを第2のスイッチングLSI群82-1～82-Mにより接続し、これらのLSI群によって、Y軸クロスバースイッチ31Cと31Dの間のZ軸方向のスイッチングを行う。

上記第1、第2のスイッチングLSI群は、同一Y座標をもつLSI:80-iと82-i (i=1～M)の間を光モジュール83、84と光ファイバ300を介して結合する。

【0064】

第1LSI群に属した各スイッチングLSI:80-iのインターフェースユニット81A、81Bでは、それぞれの入力パケット制御回路によって、Y軸クロスバースイッチ31Aまたは31Bから受信したパケットの宛先Z座標値zに応じて、z=1またはz=2の場合はインターフェースユニット81Aと81Bの間で、z=3の場合はインターフェースユニット81C宛に、z=4の場合はインターフェースユニット81D宛にパケットを転送制御する。一方、第2LSI群に属した各スイッチングLSI:82-iのインターフェースユニット81A、81Bでは、Y軸クロスバースイッチ31Cまたは31Dから受信したパケットの宛先Z座標値zに応じて、z=1の場合はインターフェースユニット81C宛に、z=2の場合はインターフェースユニット81D宛に、z=3またはz=4の場合はインターフェースユニット81A、81B間でパケットを転送制御する。

本実施例によれば、第1、第2のスイッチングLSI群80、82のパケットスイッチング機能を利用して、ノード数が $L \times M \times 4$ の分散Exchanger方式3次元クロスバーネットワークを構成できる。

【0065】

【発明の効果】

以上の説明から明かなように、本発明によれば、多次元クロスバーネットワークおよび並列計算機システムの少なくとも一部のクロスバースイッチ間に光インターフェースを備えたスイッチング装置を介在させることによって、システム性能を低下させることなく、接続可能なノード数を増加できるという効果がある。

【図面の簡単な説明】

【図1】

本発明によるExchanger方式の3次元クロスバーネットワークにおけるX、Y、Z軸クロスバースイッチ間の接続関係を説明するための図。

【図2】

演算ノードが搭載されるノードボード10の構成を示す図。

【図3】

本発明の分散Exchanger方式の3次元クロスバーネットワークにおけるX軸クロスバースイッチとY軸クロスバースイッチとの接続関係を示す図。

【図4】

本発明の分散Exchanger方式の3次元クロスバーネットワークにおけるY軸クロスバースイッチとZ軸クロスバースイッチとの接続関係を示す図。

【図5】

インターフェース変換LSI:32の構成を示すブロック図。

【図6】

インターフェース変換LSI:32の実用的な形態の1例を示す図。

【図7】

Y軸クロスバースイッチの主要部を示す図。

【図8】

Y軸クロスバースイッチの変形実施例を示す図。

【図9】

本発明によるExchanger方式の3次元クロスバーネットワークの第2の実施例を示す図。

【図10】

図9におけるスイッチングLSI:80の1実施例を示す図。

【図11】

本発明によるExchanger方式の3次元クロスバーネットワークの第3の実施例を示す図。

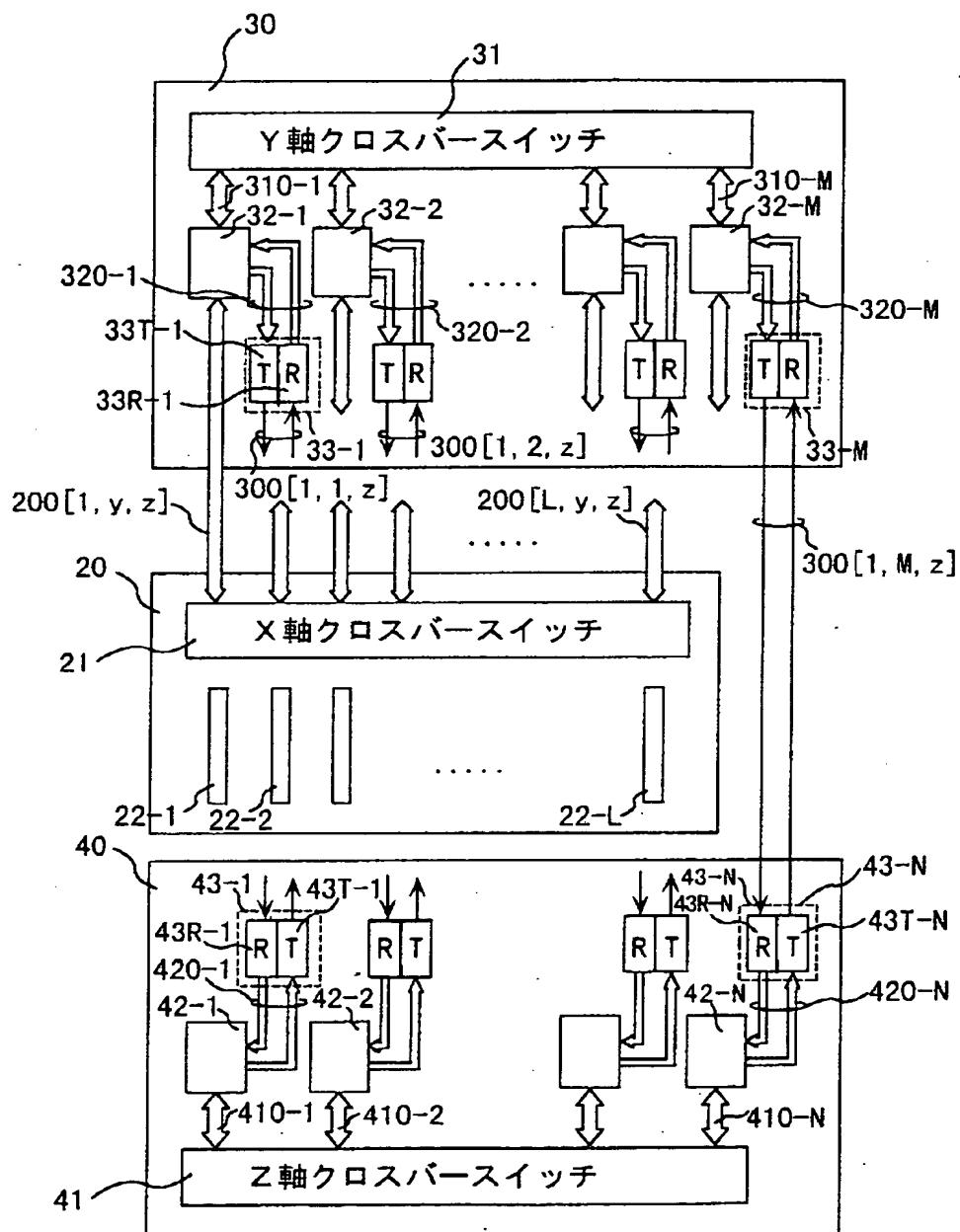
【符号の説明】

10:ノードボード、20:X軸ボード、
21:X軸クロスバースイッチ、22:ノードボード接続コネクタ、
30:Y軸ボード、31:Y軸クロスバースイッチ、
32:インターフェース変換LSI、33:光モジュール、
40:Z軸ボード、42:インターフェース変換LSI、43:光モジュール、
51:入出力ポート、52:位相調整回路、53:データバッファ、
54:入力パケット制御回路、55:セレクタ、56:出力制御回路、
57、58:同期化回路、59:データ幅変換回路、60、61:PLL回路、
200:入出力線(同軸ケーブル)、300:入出力線(光ファイバ)。

【書類名】 図面

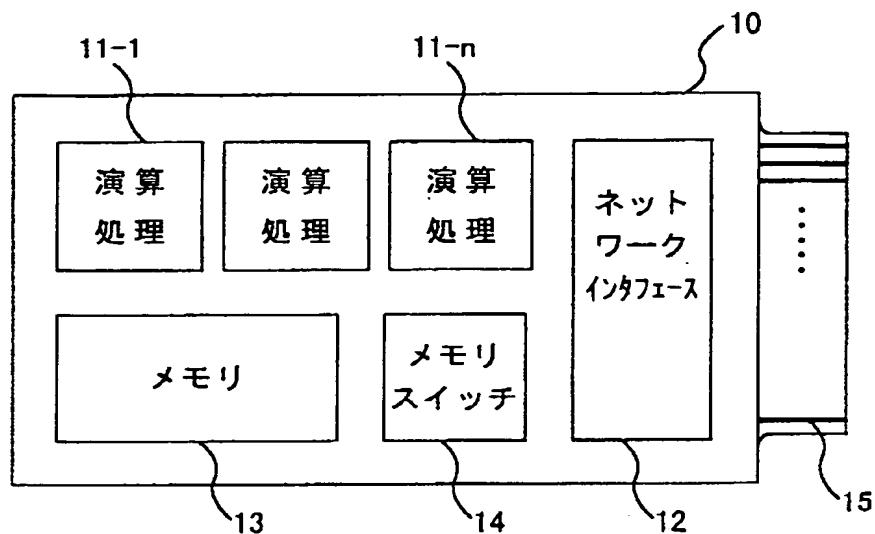
【図1】

図 1



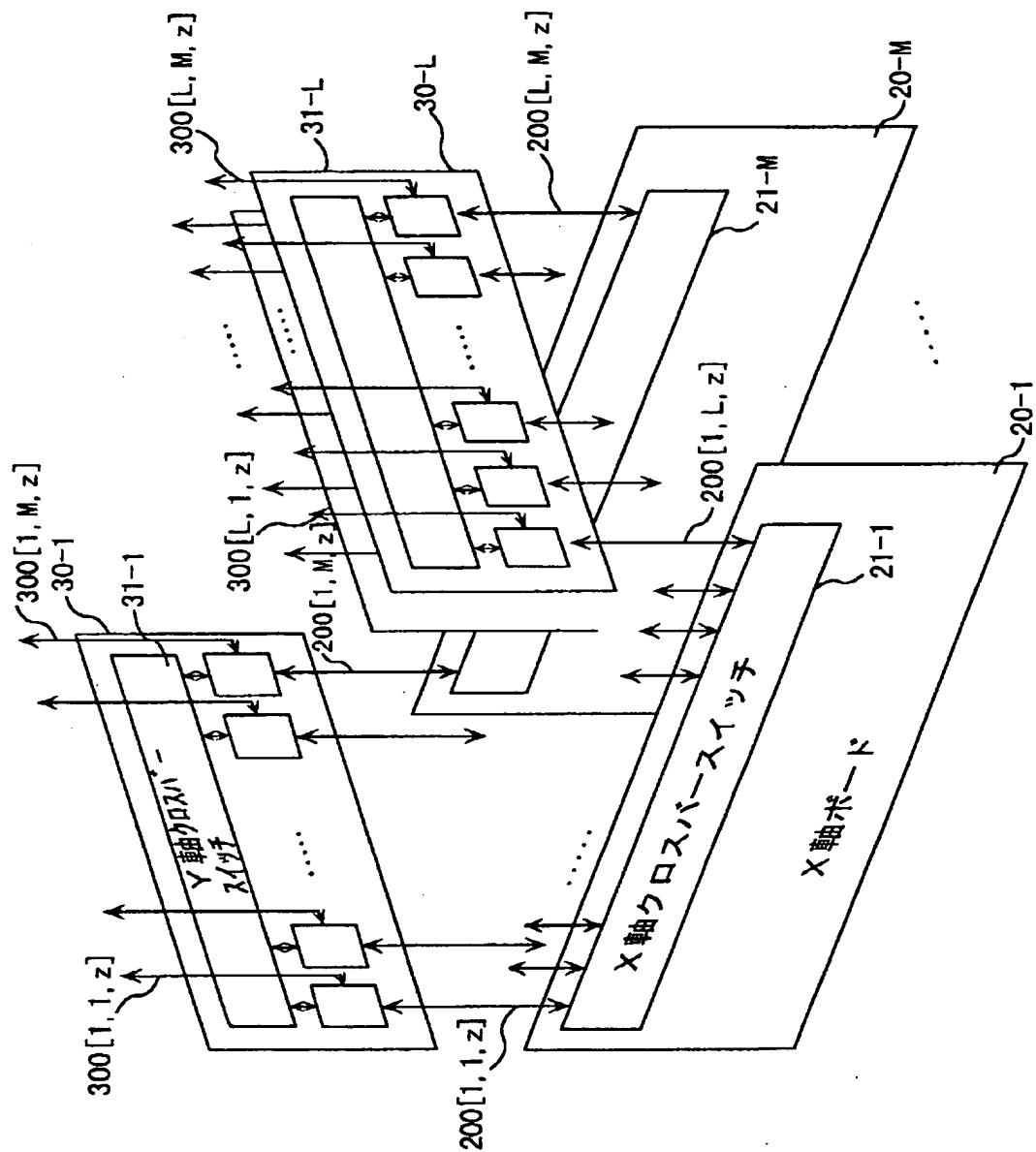
【図2】

図 2



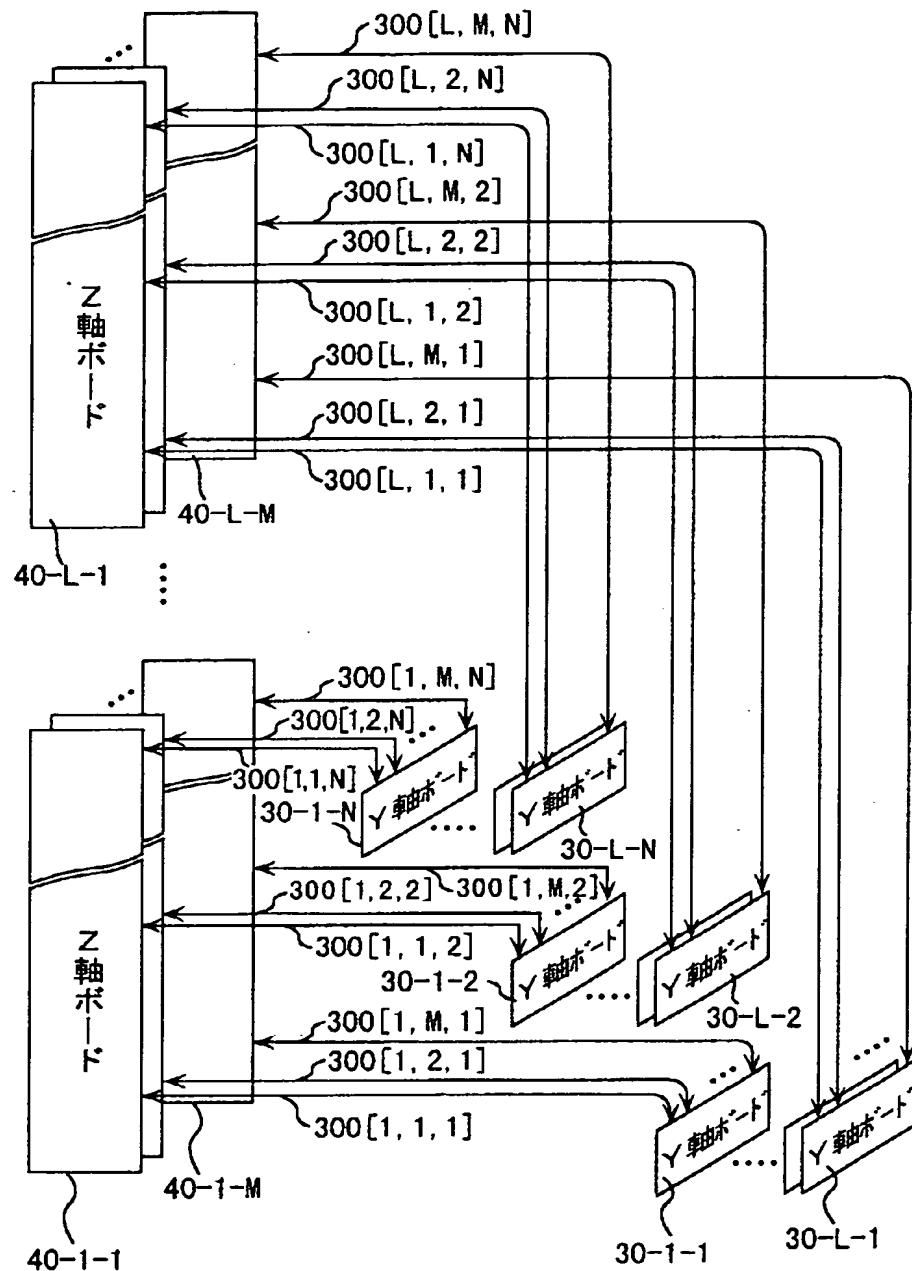
【図3】

図 3



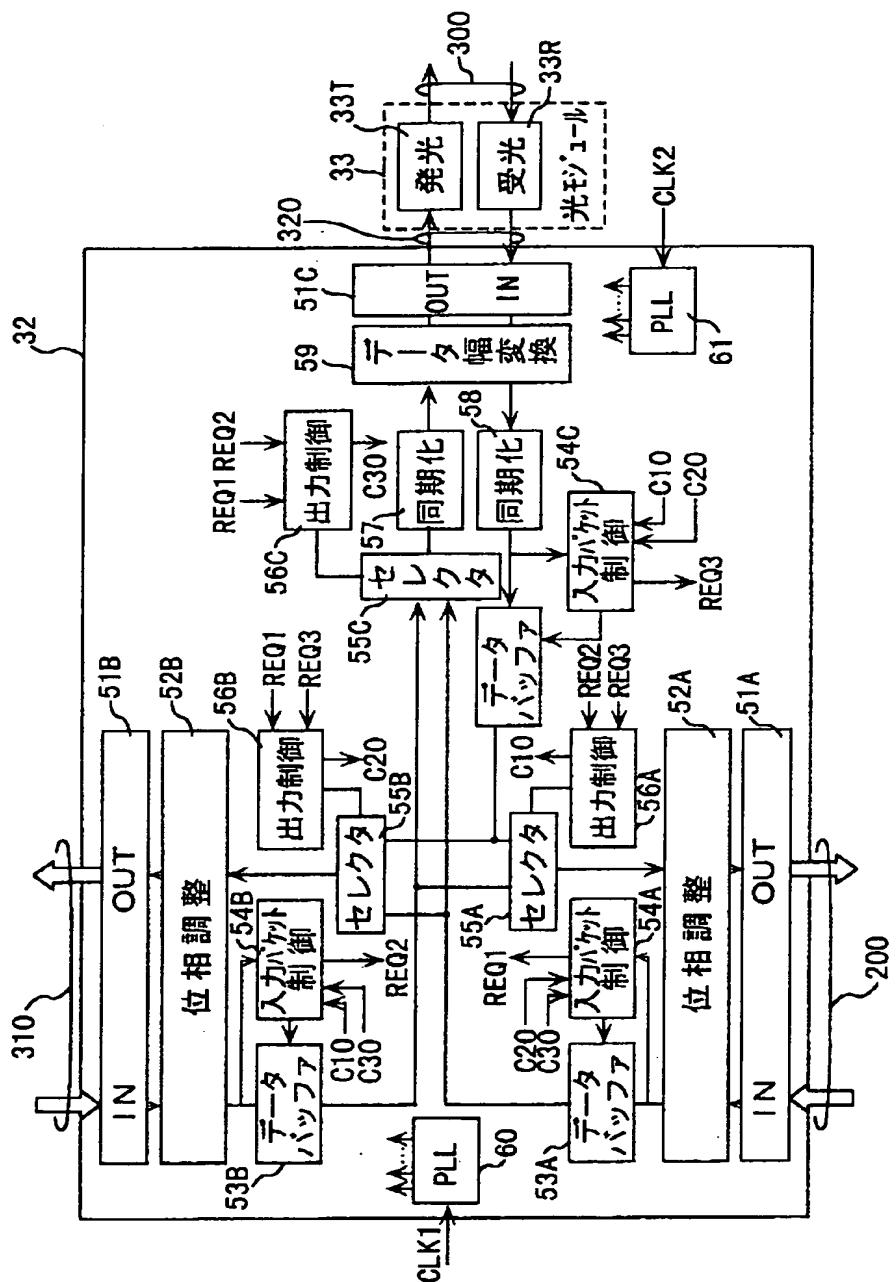
【図4】

図 4



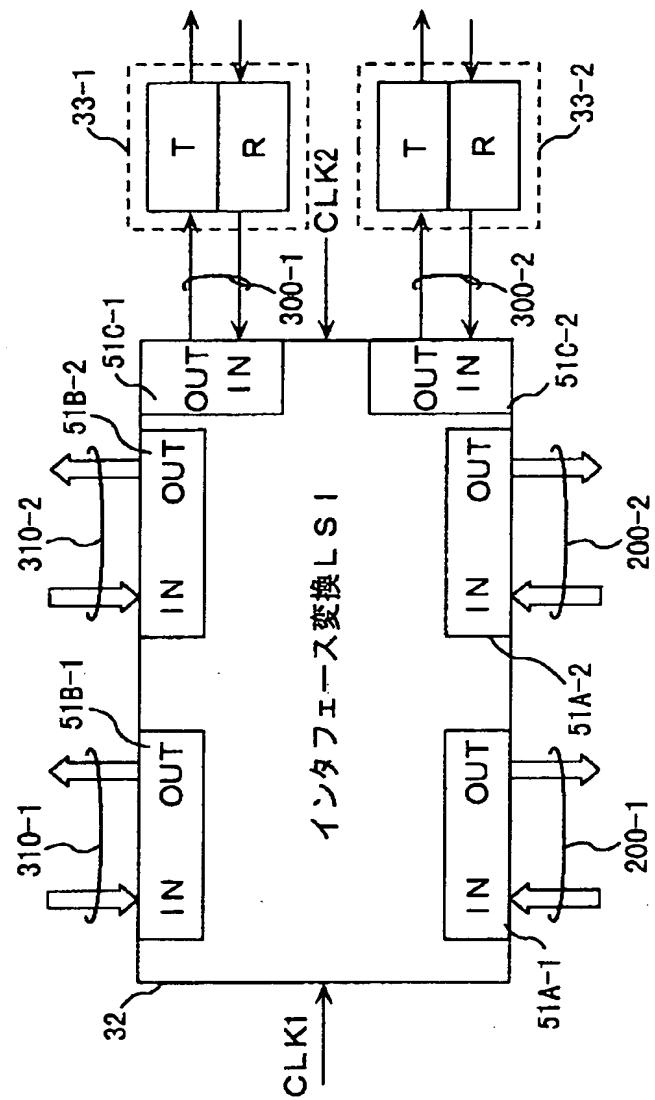
【図5】

図 5



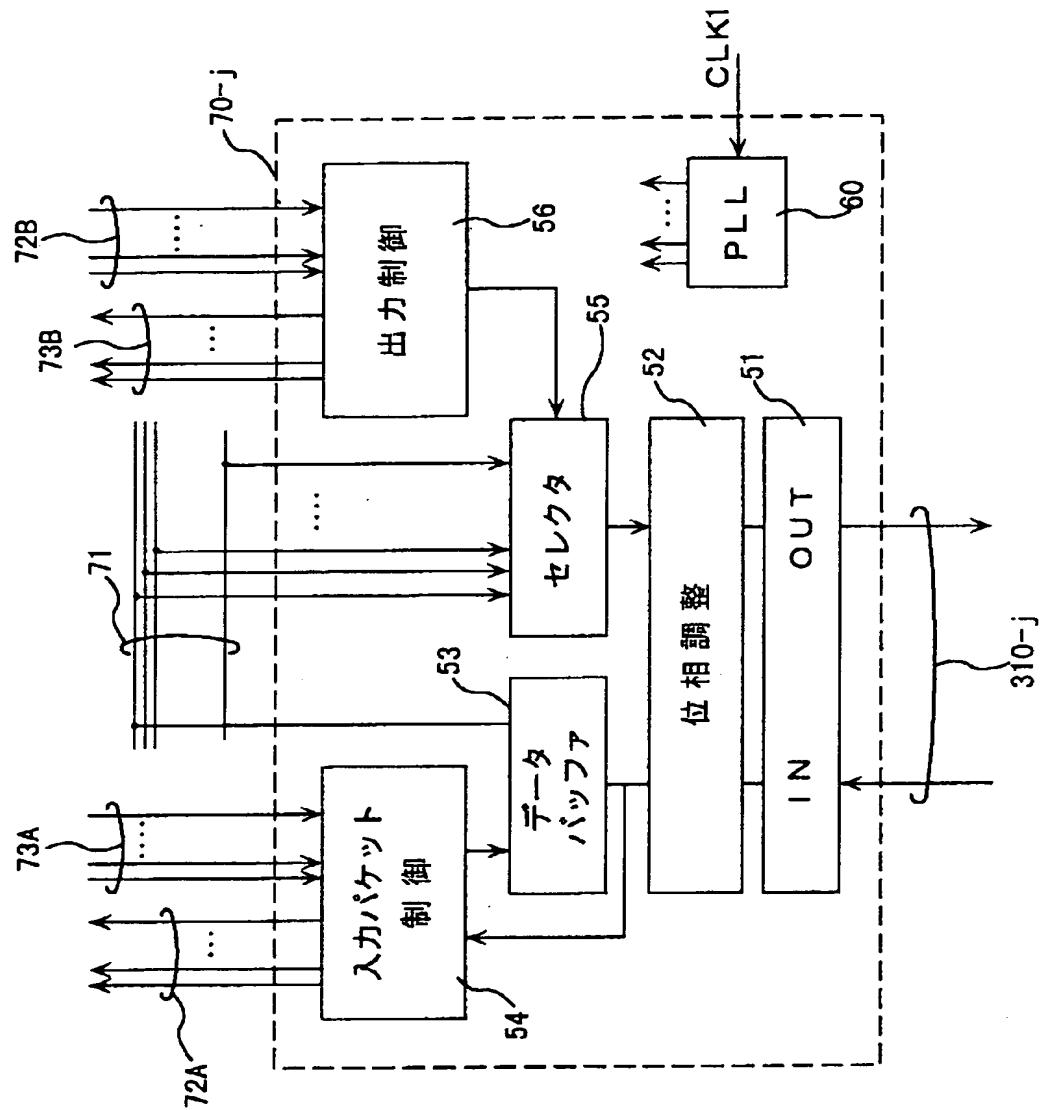
【図6】

図 6



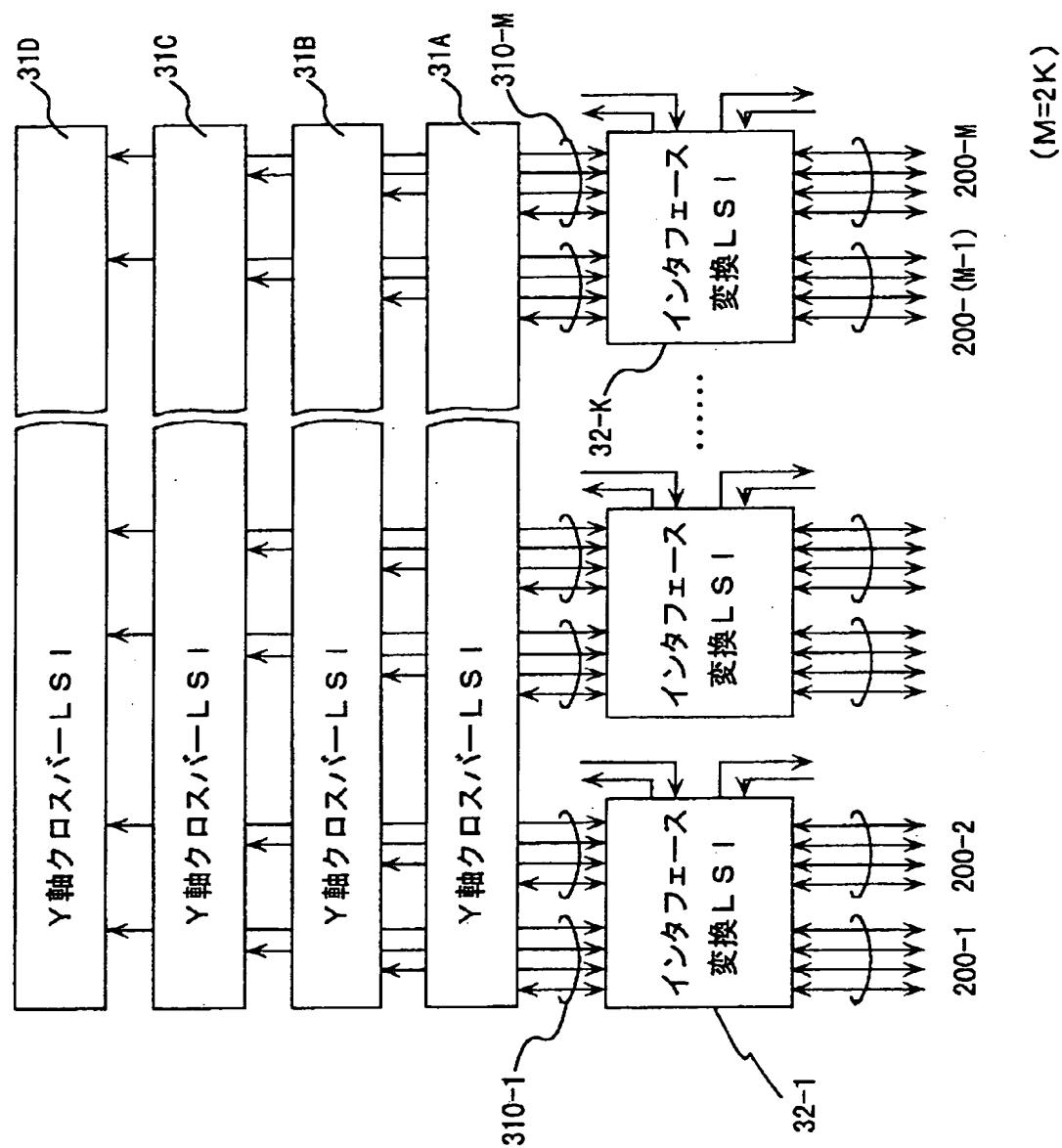
【図7】

図 7



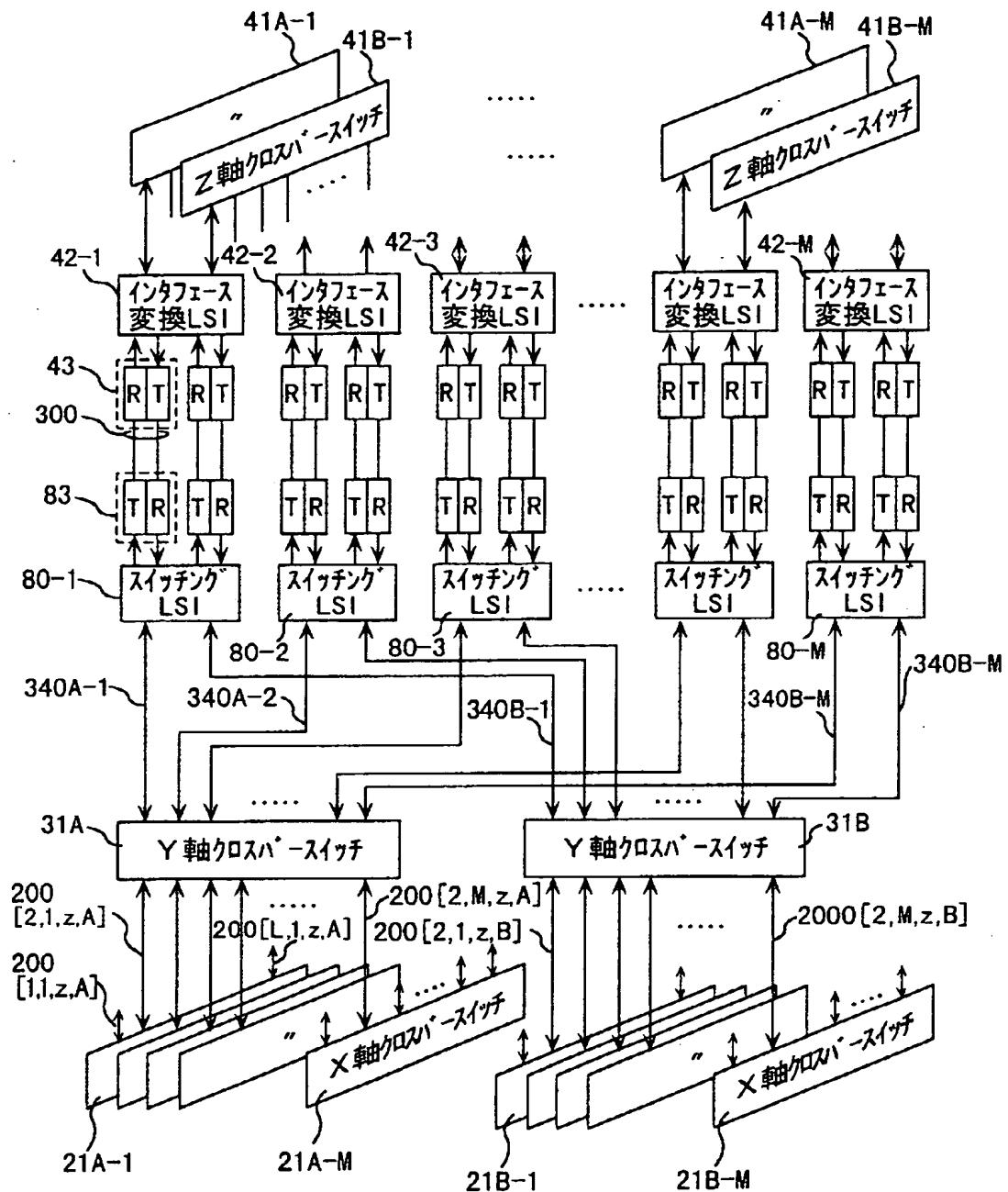
【図8】

図 8



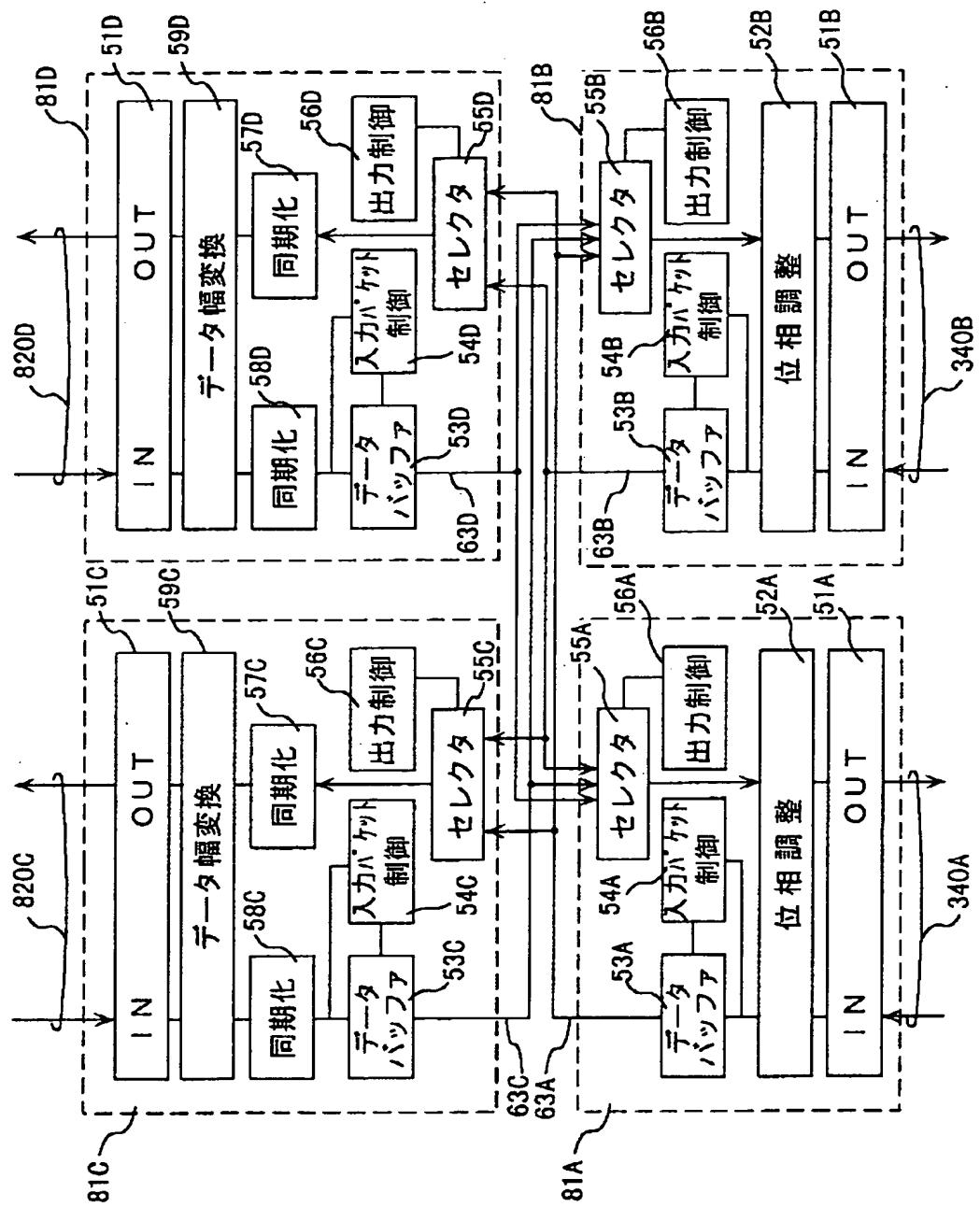
【図9】

図 9



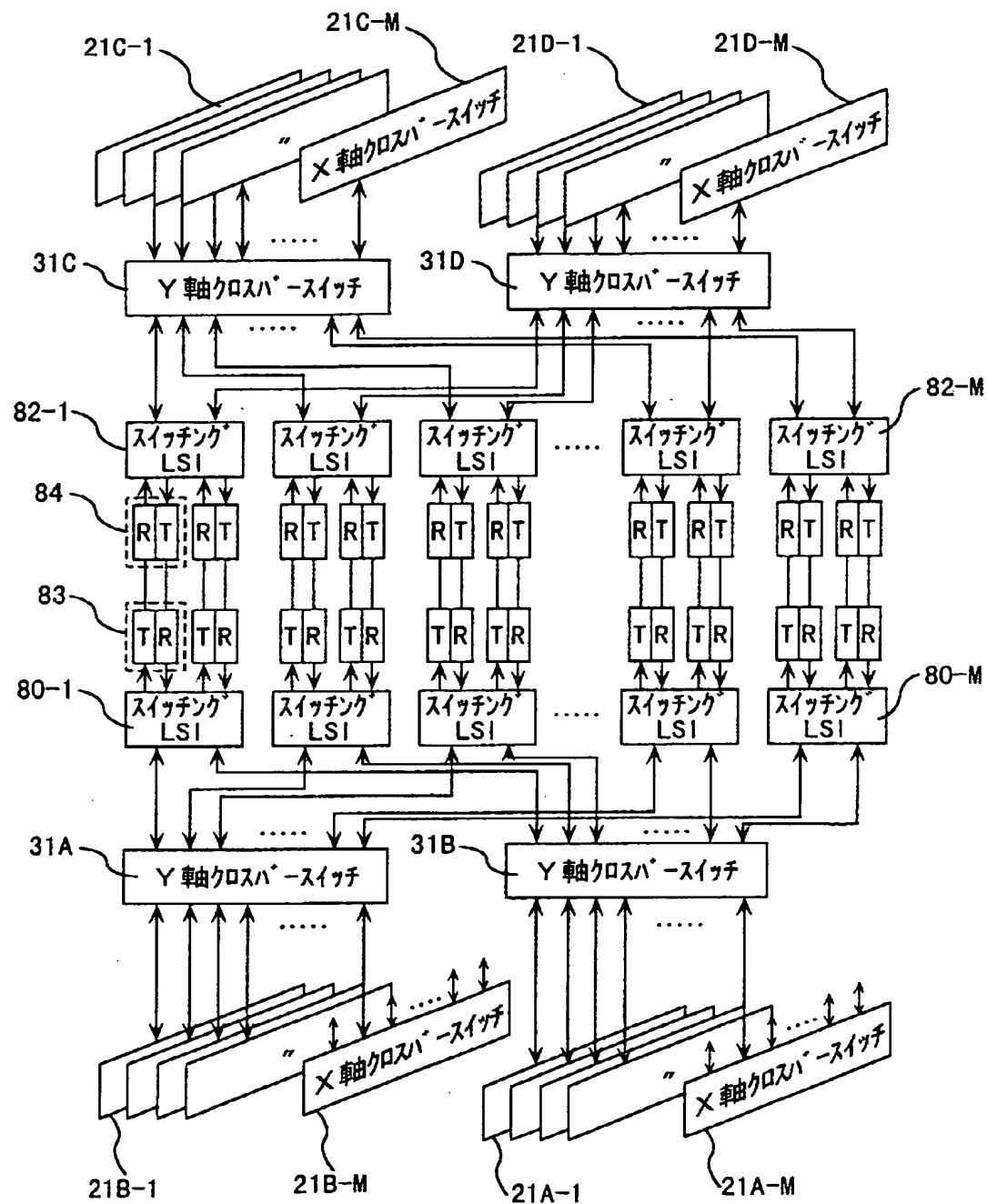
〔図10〕

図 10



【図11】

図 11



【書類名】 要約書

【要約】

【課題】 分散Exchanger方式の多次元クロスバーネットワークおよび並列計算機システムにおいて、接続可能なノード数を増加すること。

【解決手段】 $L \times M \times N$ 個の演算装置を接続する分散Exchanger方式の3次元クロスバーネットワークにおいて、X軸クロスバースイッチ21とY軸クロスバースイッチ31との間に、光モジュール33を備えたインタフェース変換LSI32を介在させ、Z軸クロスバースイッチ41へのパケット転送を上記インタフェース変換LSIで行い、Y軸、Z軸クロスバースイッチ間のパケット転送路に光ファイバ300を適用する。

【選択図】 図1

出願人履歴情報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所